CHARACTERIZATION OF HORIZONTALLY-TRANSFERRED VITAMIN B BIOSYNTHESIS GENES IN THE SOYBEAN CYST NEMATODE, *HETERODERA GLYCINES*

BY

IBTESAM NAZIM AHMED

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Master's Committee:

Associate Professor Kris Lambert, Chair
Associate Professor Leslie Domier
Professor Dean Riechers

Abstract

Heterodera glycines (*H. glycines*), commonly known as the soybean cyst nematode (SCN), is an obligate plant parasite and a major pest of soybean. It is highly adapted to manipulate and parasitize host plants. It possesses parasitism/effector genes that enable it to developmentally alter host cells and establish feeding sites inside the host root, and also suppress or counteract host defenses. Currently, there is a lot of interest in identifying and characterizing genes that enable this nematode to parasitize its host in order to find molecular targets for controlling this pest. Seven *H. glycines* genes, namely *panC*, *bioB*, *tenA*, *thi4*, *thiD*, *thiE* and *thiM*, have been discovered recently. All these genes are involved in vitamin B biosynthetic pathways forming one complete and two partial pathways. They are potentially important molecular targets because there is evidence to show that these genes have been acquired through horizontal gene transfer from prokaryotes, and have been implicated to play roles in plant parasitic nematode-host interactions.

The main goal of this project was to perform complementation experiments in order to determine the function of each of the seven SCN vitamin biosynthesis/salvage pathway genes. To achieve this goal, the cDNA of each of the genes was PCR-amplified using gene-specific primers. Each cDNA insert was cloned into a cloning vector, and transformed into chemically competent *Escherichia coli* (*E. coli*) cells. The recombinant plasmid DNA was used to transform mutant *E. coli* cells deficient in that particular gene. The growth of these transformed mutants was analyzed, in order to understand whether there was complementation. The *thiD* complementation experiment showed the most clear-cut results. A mutant strain of *E. coli* deficient in *thiD* and transformed with *thiD*-containing vector grew much better in M9 minimal media than the same strain of cells transformed with vector alone. Results from this study also indicated that *H. glycines panC*, *tenA*, *bioB*, *thiE* partially complemented their corresponding mutant *E. coli* strains.

Acknowledgements

It is with immense gratitude that I acknowledge the help of my advisor, Professor Kris Lambert. Without his guidance and continuous support, this thesis would not have been possible. I am also grateful to my committee members Professor Leslie Domier and Professor Dean Riechers, for their support and guidance throughout this project. I would like to thank Dr. Sadia Bekal for always giving me encouragement and moral support. I would also like to thank my colleague Biruk Gonfa for helping me with laboratory protocols. Finally, I would like to thank my family and friends for their encouragement and support.

Table of Contents

1 Introduction

Nematodes are a large phylum of animals that includes plant and animal parasites as well as many free-living species. The soybean cyst nematode (SCN), *Heterodera glycines* (*H. glycines*), is a small, obligate plant-parasitic roundworm that infects the roots of soybean plants. It is the most damaging, most economically-important pathogen of soybean and the leading cause of soybean yield loss in the United States (Wrather and Koenning, 2009) and other soybean-producing countries worldwide (Wrather *et al.*, 2001). Severely affected plants show symptoms such as leaf and stem chlorosis, root necrosis, reduced shoot and root growth, and loss of seed yield. The soybean cyst nematode has been estimated to cause the United States several hundred thousand tons of soybean yield losses every year (Wrather and Koenning, 2006). One reason why yield losses are so great is that growers have been cultivating soybean too intensively and frequent cropping favors the nematode. Global soybean cultivation has dramatically increased over the last few years as the demand for soybean-based products have escalated (Schmitt *et al.*, 2004).

Currently, common control practices include crop rotation with non-hosts and planting resistant soybean varieties. Planting the same SCN resistant soybean cultivar repeatedly may cause new virulent Hg Types of SCN to appear and this may reduce the ability of those soybean cultivars to control the nematode (Schmitt *et al.*, 2004). *H. glycines* parasitism is designated by Hg Type, in which parasitism is assessed based on how well SCN can reproduce on a set of seven *H. glycines*-resistant soybean lines (Niblack *et al.*, 2002). Once a field is infested it is not possible to completely eradicate this pest, mainly because the cysts remain viable for years in the soil. Nematicides are not recommended because they increase production costs and are toxic to human health and the environment. There is a pressing need for discovering specific targets and developing novel strategies to combat this pest problem. In order to come up with innovative management strategies, we need to understand the biology of this nematode in order to determine why this worm is such a successful pathogen.

If we look at the life cycle of SCN, it becomes clear that this worm is a complex and an extremely well-adapted parasite (more details in Chapter 2). Although the events of the life cycle are well-known, there is not enough knowledge about the molecular players involved in this

complex host-parasitic interaction. The nematode uses its stylet to introduce a complex mixture of proteins/metabolites into a host root cell and transforms it into a syncytium, and feeds from it for the rest of its life. Currently, there is a lot of research interest on elucidating the molecular mechanisms of host-parasitic interaction and identifying molecular targets.

Parasitism/effector genes from phytoparasitic nematodes are thought to be essential for nematode invasion of the host plant. For instance, chorismate mutase, a secreted enzyme produced in the nematode's esophageal glands, is one gene that is thought to play several roles in nematode parasitism (Lambert *et al.*, 1999). Chorismate mutase is an enzyme in the shikimate pathway which is a primary metabolic pathway found in plants, bacteria, fungi and some protozoans (Roberts *et al.*, 1998; Schmid and Amrhein, 1995). The shikimate pathway is crucial for plants because it is an important metabolic route that produces plant hormones, components of cell wall, aromatic amino acids, and various secondary metabolites that are involved with plant defense responses (Dewick, 1998). The shikimate pathway does not exist in animals, which makes the expression of a secreted form of nematode chorismate mutase by nematodes especially intriguing. Because chorismate mutase is a nematode enzyme that can potentially manipulate the plant shikimate pathway and suppress host defense compounds, it was hypothesized that it has the ability to enhance the parasitic ability of nematodes expressing the gene and enable the nematode to attack even resistant soybean plants (Bekal *et al.*, 2003; Lambert *et al.*, 2005).

In the past few years, identification of genes coding for secreted proteins was a breakthrough in phytoparasitic nematode research. It was previously thought that whatever molecular and developmental changes are happening during syncytium formation, are all due to what the nematode is secreting into the plant cell. Studies involving nematode-host interaction largely focused on identifying and characterizing genes encoding secreted parasitism-associated proteins such as enzymes that alter the cell wall such as cellulases (Smant *et al.*, 1998) and pectate lyases (Doyle and Lambert, 2002), and antioxidant enzymes such as peroxidases, superoxide dismutase (Zacheo *et al.*, 1983, Molinari and Miacola, 1997), chorismate mutase (Lambert *et al.*, 1999), and so on. Many of these proteins are closely related to bacterial proteins, leading to the hypothesis that these genes have been horizontally transferred from prokaryotes to phytoparasitic nematodes. However, this left non-secreted proteins uninvestigated. Nine SCN genes involved in

the biosynthesis and salvage of non-secreted B vitamins were discovered (Craig *et al.*, 2008, Craig *et al.*, 2009), and these are *HgSNO, HgSNZ*, *panC, bioB, thiD, thiM, thi4, tenA* and *thiE*. It was found that *HgSNO, HgSNZ* encode a functional pyridoxal 5'-phosphate (PLP) synthase that produces the active metabolite vitamin B_6 (Craig *et al.*, 2008). These two genes comprised the first example of a complete *de novo* vitamin biosynthesis pathway in SCN and they both successfully complemented a vitamin B_6 *E. coli* mutant. This was a very surprising discovery because up to that point, it was presumed that like all other animals, the soybean cyst nematode also lost its ability to synthesize vitamins. Studies of *Caenorhabditis elegans* (*C. elegans)* have shown that like all animals, it also has a strict requirement for B vitamins and other essential amino acids, which are absorbed in the intestine and salvaged. Therefore, it is quite unusual that despite being a parasite, the soybean cyst nematode possesses these vitamin biosynthesis genes. It was confirmed that these genes were part of the SCN genome and not prokaryotic contamination (Craig *et al.*, 2008). We suspect that these vitamin genes have important roles in host-nematode interactions, which is why in this project we have cloned each of these genes and performed complementation experiments to determine the functionality of these genes.

The rest of the thesis is organized as follows. Chapter 2 provides background information on the life cycle of the soybean cyst nematode and the three vitamin biosynthesis pathways the seven genes belong to. Chapter 3 describes the cloning and complementation experiments performed for each gene. This chapter explains the results and provides the interpretations of the results. Chapter 3 also provides details about how the experiments are set up, what reagents and equipment are used and where the materials are obtained from. Chapter 4 discusses the findings in this project, and provides directions for future works. Lastly, Chapter 5 concludes the thesis.

2 Background

2.1 Life cycle of the soybean cyst nematode

The life cycle of SCN begins at the egg stage. Embryogenesis leads to the development of the first-stage juvenile which remains inside the egg. The first-stage juvenile molts inside the egg to form a second-stage juvenile (J2). The J2 that hatches from the egg moves through the soil, searching for suitable host plants. The J2 must successfully find a host within a short time, before its energy reserves are depleted. J2s are able to sense chemicals that leach out from host plants into soil and pH gradients (Craig, 2009). Once a host is located, the J2 penetrates the root and migrates through the root tissue until it comes close to vascular bundles. It then initiates the formation of a feeding site or syncytium from which it will feed from for the rest of its life.

The juvenile uses its stylet to pierce plant cell walls and to secrete a complex mixture of metabolites and enzymes like cellulases (Smant *et al.*, 1998), pectate lyases (Doyle and Lambert, 2002) and expansins (Qin *et al.*, 2004) that help break down the plant cell walls. The secretions come from three esophageal gland cells in the nematode. The injected cell undergoes a dramatic transformation and starts incorporating surrounding cells to become a large multinucleate syncytium (Grundle *et al.*, 2004). The syncytium maintains high metabolic activity and becomes dense with organelles. At the border between the phloem and the syncytium, the cell wall forms invaginations that increase the surface area for greater absorption of nutrients. Once the feeding site is established, the nematode becomes sedentary. The nematode feeds, grows and molts through the third and fourth stages until it reaches adulthood. When males reach adulthood, they regain their vermiform shape and leave the root, looking for females to mate with. As the female nematode matures, its body expands and attains a lemon-shape. After fertilization, the female starts producing hundreds of eggs. Most eggs are retained within the body of the female but some are released into a gelatinous matrix secreted by the female. At some point, the swollen female bodies become so large that they burst out from the root epidermis. When the female dies, its body hardens and turns into a protective cyst. The eggs can remain protected inside the cyst for years, which makes SCN such a tenacious pest (Reviewed by Niblack *et al.*, 2006).

2.2 Pantothenate

Pantothenate is an essential metabolite for all biological organisms. Pantothenic acid (vitamin B5) is the essential precursor to coenzyme A (CoA) which is an important cofactor in the citric acid cycle and also in fatty acid and cholesterol biosynthesis. The pathway comprises four enzymatic steps in *E. coli* and pantothenate is thought to be synthesized in plants via a similar route (Reviewed by Webb *et al*., 2004). Pantothenate is synthesized by the condensation of pantoate and β-alanine. In *H. glycines*, the SCN homolog of *panC* gene has been discovered which means that only the partial pathway for pantothenate biosynthesis exists in this nematode (Craig *et al.*, 2009). The enzyme pantothenate synthetase, encoded by *panC*, catalyzes the final step of the pantothenate biosynthesis pathway. In this step, pantothenate is synthesized by condensation of pantoate and β-alanine (Genschel *et al.*, 1999).

2.3 Thiamine

Thiamine (vitamin B1) plays important roles in many aspects of carbohydrate and amino acid metabolism. The thiazole and pyrimidine moieties of thiamine monophosphate are synthesized in separate branches of the pathway. Thiamine monophosphate is phosphorylated to generate thiamine diphosphate or pyrophophate, the active cofactor (Begley *et al*., 1999).

Thiamine pyrophosphate is a cofactor for many enzymatic reactions. For instance, thiamine pyrophosphate activates decarboxylation of pyruvate in the pyruvate dehydrogenase complex of the citric acid cycle. The pyruvate dehydrogenase complex is a combination of various enzymes and cofactors and pyruvate is derived from glucose (Harris *et al.*, 2002). Therefore, oxidation of glucose is highly dependent on thiamine diphosphate.

Twelve genes involved in thiamin biosynthesis have been identified in prokaryotes (Reviewed by Begley *et al.*, 1999). Six of these genes are required for the thiazole biosynthesis (*thiF*, *thiS*, *thiG*, *thiH*, *thiI*, and *dxs*) (Park *et al.*, 2003), one is involved in the pyrimidine biosynthesis (*thiC*), one is required for the linking of the thiazole and the pyrimidine (*thiE*), and four kinase genes (*thiD*, *thiM*, *thiL*, and *pdxK*) carry out phosphorylation reactions in the pathway (Reviewed by Jurgenson *et al*., 2009). ThiM enzyme is involved in the salvage of thiazole. The genes *tenA*, *thiD*, *thiE*, *thiM* and *thi4* have been discovered in *H. glycines*, potentially forming a complete pathway for *de novo* thiamine biosynthesis (Craig *et al.*, 2009). In *Bacillus subtilis*, 4-amino-2-

methyl-5-hydroxymethylpyrimidine (HMP) biosynthetic gene is *tenA*, which has been cloned and sequenced and shows high sequence similarity to the *thiC* gene in *E. coli* (Zhang *et al.*, 1997). ThiD is a HMP kinase and catalyzes the phosphorylation of HMP to HMP-P and HMP-PP and it is essential for the synthesis of thiamin. Thiamin phosphate synthase catalyzes the coupling of the pyrimidine pyrophosphate and the thiazole phosphate to give thiamin phosphate. The final phosphorylation step is catalyzed by thiamin phosphate kinase (encoded by *thiL*), which gives thiamin pyrophosphate, the biologically active form of the vitamin (Webb and Downs, 1997).

2.4 Biotin

Biotin (vitamin B7), first identified as a yeast growth factor, is an essential cofactor for biotin-dependent enzymes that catalyze carboxylation and decarboxylation reactions, in which biotin serves as a site for carbon dioxide attachment (Knowles, 1989). Biotin-dependent enzymes participate in important metabolic processes such as gluconeogenesis, fatty acid biosynthesis, and amino acid metabolism (Jitrapakdee and Wallace, 2003). The most common biotin-dependent enzyme is acetyl-CoA carboxylase, which catalyzes the ATP-dependent transfer of a carboxyl group from carbonate to acetyl-CoA to form malonyl-CoA. In *E. coli*, the genome encodes a biosynthetic pathway that produces biotin from pimeloyl-CoA in four enzymatic steps (Reviewed by Streit and Entcheva, 2003). In *H. glycines*, the SCN homolog of *bioB* gene has been discovered (Craig *et al.*, 2009) which appears to encodes biotin synthase (BioB). The BioB enzyme catalyzes the last step of the pathway. Thus, *H. glycines* is only equipped with the partial biotin biosynthesis pathway.

3 Experiments and Results

This chapter describes what experimental steps were taken during the cloning and complementation testing of each gene. The observations made during the experiments and the results obtained are reported.

3.1 Cloning of *bioB* cDNA and Δ*bioB* complementation experiment

HgBioB cDNA was amplified using the designed primers for *bioB*. A pre-constructed plasmid containing the *HgBioB* insert served as the DNA template of the PCR reaction. The *HgbioB* amplicon was gel-purified and then cloned into the linearized, gel-purified pBAD.LIC.A8 plasmid vector, forming the pBAD.LIC.A8.HgBioB plasmid. This plasmid was then transformed into chemically competent *E. coli* TOP10 cells. Transformed cells were selected on LB, ampicillin (100 mg/mL) plates. Next, colonies picked from these selective LB, ampicillin plates were grown for 12-16 hours in liquid LB media containing ampicillin. Plasmid DNA was extracted from these overnight cultures, and transformed into kanamycin resistant, electrocompetent Δ*bioB E. coli* mutant. Transformed cells were selected on LB, ampicillin, kanamycin plates. Δ*bioB* cells were also transformed with empty pBAD vector, and these were used later as negative control in the complementation experiment. The pBAD.LIC.A8.HgBioB plasmid DNA was sent to the sequencing center to verify whether the insert had the correct sequence.

The transformed mutant Δ*bioB E. coli* cells were incubated overnight in the shaking incubator at 200 rpm and 37˚C, in LB medium containing both ampicillin and kanamycin antibiotics. These cells are streaked on to M9, avidin alone (+/- arabinose), and M9, avidin and biotin plates (+/- arabinose). It seemed that there was some growth in the M9, avidin alone plates, hinting that there may be some partial complementation. The sequencing result showed that the BIOB insert had some mutations in it. Therefore, new colonies were picked from the LIC-cloning LB, ampicillin plates, and plasmid DNA was extracted. These plasmid DNA samples were sent again for sequencing, and this time the insert sequence had no mutations. Mutant *E. coli* cells deficient in *bioB* were transformed again with pBAD.LIC.A8.BioB and empty pBAD.LIC.A8 vector. The transformed mutant cells were plated on M9, avidin and M9, avidin, biotin plates. After 24 hours, there was no difference between the positive control and negative control plates.

However, when they plates were checked again after about two weeks, a few large colonies were found in the plates that were streaked with Δ*bioB E. coli* transformed with pBAD.LIC.A8.*bioB*. These colonies were grown in LB, ampicillin, kanamycin culture and plasmid DNA was extracted from them. The plasmid DNA was sent for sequencing, and the BIOB sequence was correct for all of them.

In order to show the difference between the growth rates of mutant cells transformed with pBAD.LIC.A8.BIOB versus mutant cells transformed with empty pBAD vector, a growth curve experiment was performed several times. However, there was no difference in the growth rates, and all cell lines are growing profusely. It is suspected that the cell cultures got contaminated. Therefore, no reliable results could be obtained.

3.2 Cloning of *panC* cDNA and Δ*panC* complementation experiment

Firstly, *HgpanC* cDNA was PCR-amplified from a pre-constructed *panC*-cDNA-containing plasmid, using the PANC-LIC forward and reverse primers. The PCR amplicon was run on the gel to verify its size. The gel picture indicated that the size of the band was within the expected range. Therefore, the amplicon was run again on the gel and gel-purification was performed.

Next, LIC reaction was carried out for the gel-purified *panC*-amplicon and the EcoRV-digested, gel-purified pBAD.LIC.A8 vector. The LICed *panC* insert and LICed pBAD vector were mixed to let them anneal and form pBAD.LIC.A8.panC, which was then transformed into chemically competent, *E. coli* TOP10 cells. These transformed cells were plated on LB, ampicillin plates and incubated overnight at 37˚C. As expected, the cells transformed with pBAD.LIC.A8.panC grew into colonies on the selective plates and cells transformed with vector alone did not grow. Some of the colonies were picked and grown overnight in liquid LB media and plasmid DNA extraction was performed. The *panC* cDNA was amplified from these extracted plasmids and the amplicons were run on 1% agarose gel for verification. The gel picture showed that all the bands were identical and of the expected size, implying that cloning was successful and the plasmids contained the PANC insert in them. Mutant Δ*panC* cells were made electrocompetent and later transformed with pBAD.LIC.A8.panC and also empty pBAD.LIC.A8 vector. The cells were spread on LB, ampicillin, kanamycin plates and incubated at 37˚C overnight. The transformation was successful and cells grew to form bacterial lawns on the plates. There was no growth in the

negative control plate that had the untransformed mutant cells streaked on it. This indicated that the ampicillin and kanamycin used were effective in selecting cells that had the proper plasmid in them.

To avoid having to transfer chunks of these cells to M9 plates for the complementation experiment, they were regrown overnight in 5 ml liquid LB culture. On the following day, one loop of cell culture was used to inoculate each plate. This was done to maintain consistency among the plates. The following eight plates were prepared:

1) M9 plate, Δ*panC* cells with empty vector
2) M9 plate, Δ*panC* cells with empty vector + arabinose
3) M9 plate, Δ*panC* cells with pBAD.LIC.A8.panC
4) M9 plate, Δ*panC* cells with pBAD.LIC.A8.panC + arabinose
5) M9 + pantothenate plate, Δ*panC* cells with empty vector
6) M9 + pantothenate plate, Δ*panC* cells with empty vector + arabinose
7) M9 + pantothenate plate, Δ*panC* cells with pBAD.LIC.A8.panC
8) M9 + pantothenate plate, Δ*panC* cells with pBAD.LIC.A8.panC + arabinose

The M9 plates were incubated overnight at 37˚C and inspected after 24 hours (the next day). As expected, there was clear growth in all the positive control (M9 + pantothenate) plates. But there was not much difference between the negative control plates and plates with pBAD.panC cells (some slight growth in these plates). Therefore, it was difficult to assess whether there was any complementation. The complementation experiment was repeated and the second time it was done, some changes were made to the steps. The 5 ml LB overnight cultures were centrifuged and the supernatant was removed. Then 5 ml M9 media was added and the pellet of cells was resuspended in it. The cell suspension was centrifuged again and the supernatant was removed. Finally, the pellet was resuspended in 1 ml M9 media. These steps were done to wash off LB, thus removing any trace of pantothenate that may be present. Once the M9 plates were inoculated with the cells, the plates were kept at 28˚C instead of 37˚C to make conditions more optimum for nematode gene expression, since it is known that optimum nematode development occurs at 23-28°C (Riggs, 1982; Burrows and Stone, 1985). There was no growth in the negative control plate and some growth in the M9, pBAD.LIC.A8.panC-cells plate, indicating that there was some partial complementation. The pBAD.LIC.A8.panC plasmid DNA that was used to

transform the mutant Δ*panC* cells was sent for sequencing for verification. The sequencing results showed that there were no mutations in the *panC* insert.

A growth curve experiment was performed to assess the difference in growth rates in M9 minimal media between mutant cells transformed with pBAD.LIC.A8.panC and mutant cells transformed with empty pBAD.LIC.A8 vector. Mutant cell lines were grown overnight in LB media so that the cells were fresh and of the same health. Four tubes containing 5 ml M9 media were prepared and in two of them 50 µl 2% arabinose solution was added. Since arabinose is an inducer, we wanted to see whether overexpression of the plasmid has any effects on the growth rate. A single loop of cells from the LB cell culture was used to inoculate each tube of M9 media to maintain consistency among the tubes. The cell cultures were allowed to grow in the 28ºC shaking incubator and OD_{600nm} readings were taken at different time points using the spectrophotometer. Figure 1 plots the OD_{600nm} readings data from the growth curve experiment.

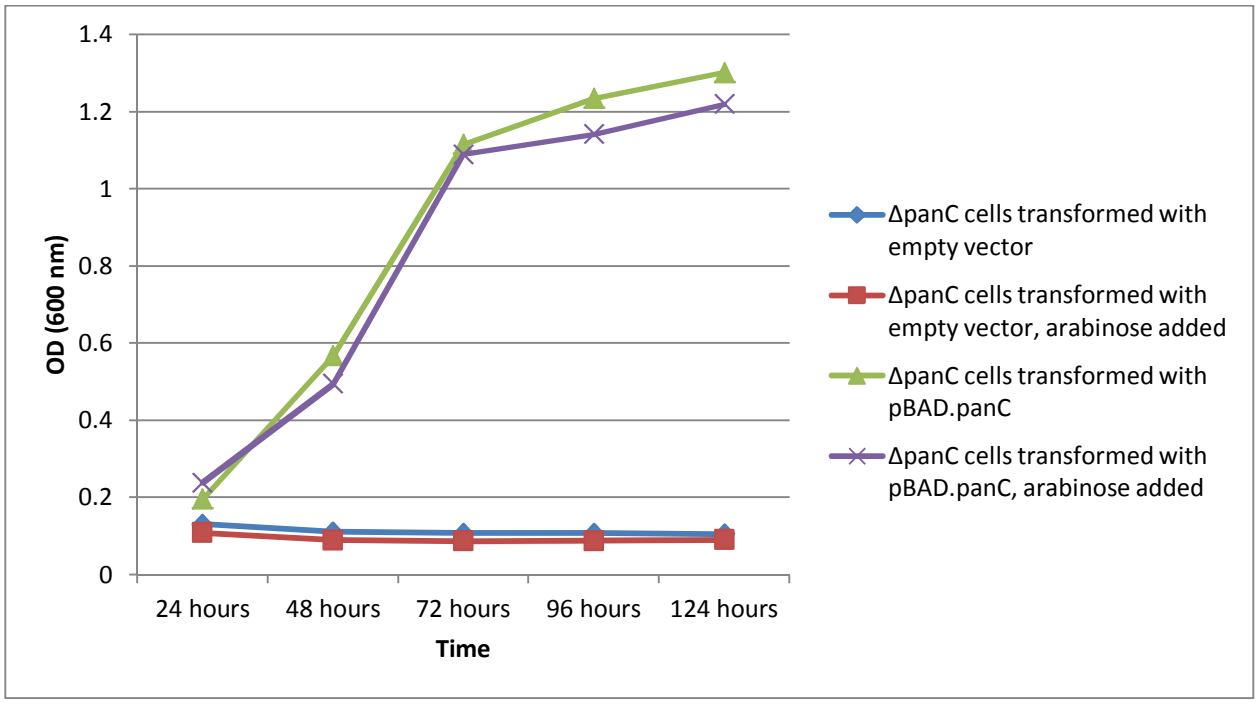


Figure 1: Growth curve analysis of Δ*panC* mutant cells in M9 media. This experiment was performed to assess the difference in growth rates in M9 minimal media between a mutant Δ*panC E. coli* strain transformed with pBAD.LIC.A8.panC and the same strain transformed with pBAD.LIC.A8 empty vector. Four tubes containing M9 media were prepared and in two of these, 50 µl of 2% arabinose solution was added to verify whether overexpression of the plasmid has any effects on the growth rate. The cell cultures were incubated at 28ºC in a shaking incubator and OD_{600nm} readings were taken at different time points using the spectrophotometer.

From Figure 1, it is clear that cells transformed with pBAD.LIC.A8.panC are growing more successfully in minimal media. When the growth curve experiment was repeated, similar results were found. This confirmed that SCN's *panC* gene was able to complement the *E. coli* mutant deficient in this gene.

3.3 Cloning of *thiD* cDNA and Δ*thiD* complementation experiment

After PCR-amplifying *thiD* cDNA by PCR using forward and reverse *thiD* primers, the amplicon was run on 1% agarose gel and gel-purification was performed. Next, LIC-cloning of gel-purified *thiD* insert and gel-purified pBAD vector was performed. The LICed insert and LICed vector were mixed and allowed to anneal and form pBAD.LIC.A8.thiD. The mixture was transformed into chemically-competent TOP10 cells. Transformation was successful and numerous colonies formed in the selective LB, ampicillin plates. Four colonies were picked and grown in LB, ampicillin media overnight in the 37˚C shaking incubator. After overnight incubation, the cells were centrifuged and plasmid DNA extraction was performed. Next, the plasmid DNA was amplified by PCR and the amplicons were run on the gel to verify whether they contain the insert. All four lanes showed bands of the right size, which showed that LIC-cloning was successful. The plasmid DNA was sent for sequencing to verify whether the insert sequence was correct. The sequencing result showed that sequence had one mutation but it was a minor change that did not affect the protein sequence. Electrocompetent Δ*thiD* mutant *E. coli* cells were transformed with pBAD.LIC.A8.thiD and also empty pBAD vector and these cells were plated on LB, ampicillin, kanamycin plates. Next, two batches of M9 minimal media plates were made, one batch containing 120μM thiamine and the required antibiotics and the other only M9 and antibiotics. Like all the other complementation experiments, the following eight plates were prepared for the Δ*thiD* complementation experiment:

1) M9 plate, Δ*thiD* cells with empty vector
2) M9 plate, Δ*thiD* cells with empty vector + arabinose
3) M9 plate, Δ*thiD* cells with pBAD.LIC.A8.thiD
4) M9 plate, Δ*thiD* cells with pBAD.LIC.A8.thiD + arabinose
5) M9 + Thiamine plate, Δ*thiD* cells with empty vector
6) M9 + Thiamine plate, Δ*thiD* cells with empty vector + arabinose
7) M9 + Thiamine plate, Δ*thiD* cells with pBAD.LIC.A8.thiD

8) M9 + Thiamine plate, Δ*thiD* cells with pBAD.LIC.A8.thiD + arabinose

These plates were incubated overnight at 28˚C. As expected, there was growth of cells in all the positive control plates (i.e., M9 + thiamine plates) and no growth in the empty control plates. Most importantly, there was also similar growth in the M9 plates that had cells transformed with pBAD.LIC.A8.thiD. Thus, full complementation was achieved.

In order to assess the differences in the rate of growth in M9 media, a similar approach was adopted as described in Section 4.2. Figure 2 plots the OD_{600nm} readings obtained from the growth curve experiment.

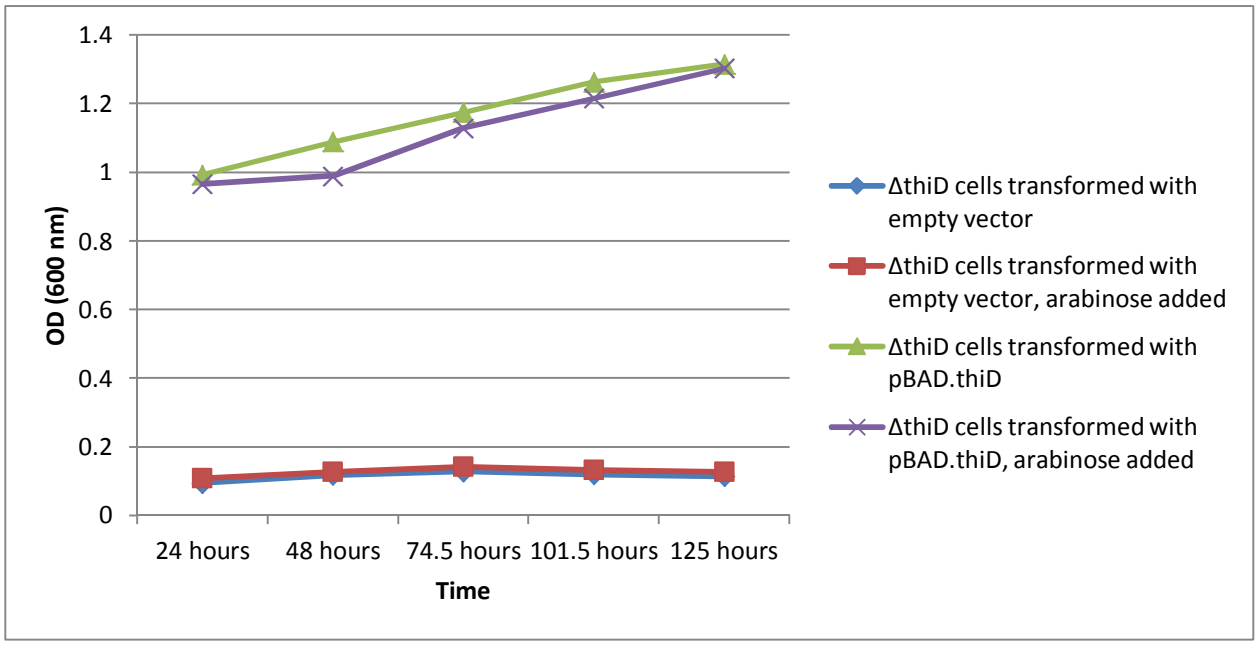


Figure 2: Growth curve analysis of Δ*thiD* mutant cells in M9 media. This experiment was performed to assess the difference in growth rates in M9 minimal media between a mutant strain of *E. coli* cells (designated Δ*thiD*) transformed with pBAD.LIC.A8.thiD, and the same strain of mutant *E. coli* cells transformed with empty pBAD.LIC.A8 vector. Four tubes containing 5 ml M9 media were each inoculated with a loop of mutant cells. In two tubes, 50 µl of 2% arabinose solution was also added in order to verify whether overexpression of the plasmid had any effects on the growth rate. The cell cultures were incubated at 28ºC in a shaking incubator and OD_{600nm} readings were taken at different time points using the spectrophotometer.

From Figure 2, it was clear that cells transformed with pBAD.LIC.A8.thiD were growing successfully in M9 minimal media and cells transformed with empty pBAD vector were not able to grow in M9 minimal media. The growth curve experiment was repeated, similar results were

found. This confirmed that SCN's *thiD* gene is able to complement the corresponding *E. coli* mutant.

3.4 Cloning of *tenA* cDNA and Δ*thiC* complementation experiment

The *tenA* cDNA was PCR-amplified using iProof master mix and *tenA* primers. The amplicons were run on the gel and gel-purified. The gel-purification was successful and bands of expected size showed up on the gel. Gel-purified *tenA* cDNA was LIC-cloned into pBAD vector to form pBAD.LIC.A8.tenA, which was used to transform chemically-competent TOP10 *E. coli* cells. These transformed cells were plated on selective LB, ampicillin plates. LIC-cloning was successful and colonies formed on the selective plates (no colonies grew on the negative control plates). Ten colonies were picked and grown overnight in 5 ml LB, ampicillin media. Plasmid DNA extraction was performed and PCR was carried for all these extracted plasmids, using *tenA* primers. The amplicons were run again for verification. The gel showed that LIC-Cloning of *tenA* was successful and all extracted plasmids contained the insert. Empty pBAD vector and pBAD.LIC.A8.tenA were used to transform electrocompetent Δ*tenA* mutant *E. coli* cells. These transformed cells were selected on LB, ampicillin, kanamycin plates.

For the complementation experiment, the following eight plates were prepared:

1) M9 plate, Δ*tenA* cells with empty vector
2) M9 plate, Δ*tenA* cells with empty vector + arabinose
3) M9 plate, Δ*tenA* cells with pBAD.LIC.A8.tenA
4) M9 plate, Δ*tenA* cells with pBAD.LIC.A8.tenA + arabinose
5) M9 + Thiamine plate, Δ*tenA* cells with empty vector
6) M9 + Thiamine plate, Δ*tenA* cells with empty vector + arabinose
7) M9 + Thiamine plate, Δ*tenA* cells with pBAD.LIC.A8.tenA
8) M9 + Thiamine plate, Δ*tenA* cells with pBAD.LIC.A8.tenA + arabinose

The plates were incubated at 28˚C overnight. From the appearance of the plates, it seemed that there was no evidence of complementation. More colonies were picked from the LIC-cloning experiment plates (LB, ampicillin). Individual colonies were grown in LB, Amp overnight culture and plasmid DNA was extracted. PCR reaction was set up in order to amplify *tenA* insert from the extracted plasmids. Moreover, plasmid DNA concentration was determined using the

spectrophotometer and the samples that had a high enough concentration of *tenA* were sent for sequencing. The sequencing result confirmed that the *tenA* insert sequence was correct (and no mutations were introduced during the cloning process).

Since the plates did not show any evidence of complementation, a growth curve experiment was conducted to verify whether there is a difference in the growth rates. The transformed cells were inoculated into M9 media and OD_{600nm} readings were taken after 24 hours, 48 hours and again after 115 hours. From the data, it was evident that cells transformed with pBAD.tenA grew better than cells transformed with empty pBAD vector. The experiment was repeated two more times and similar results were obtained. Therefore, it was concluded that *tenA* was successful in complementing its corresponding *E. coli* Δ*thiC* mutant. Figure 3 displays the data from the *tenA* growth curve experiment.

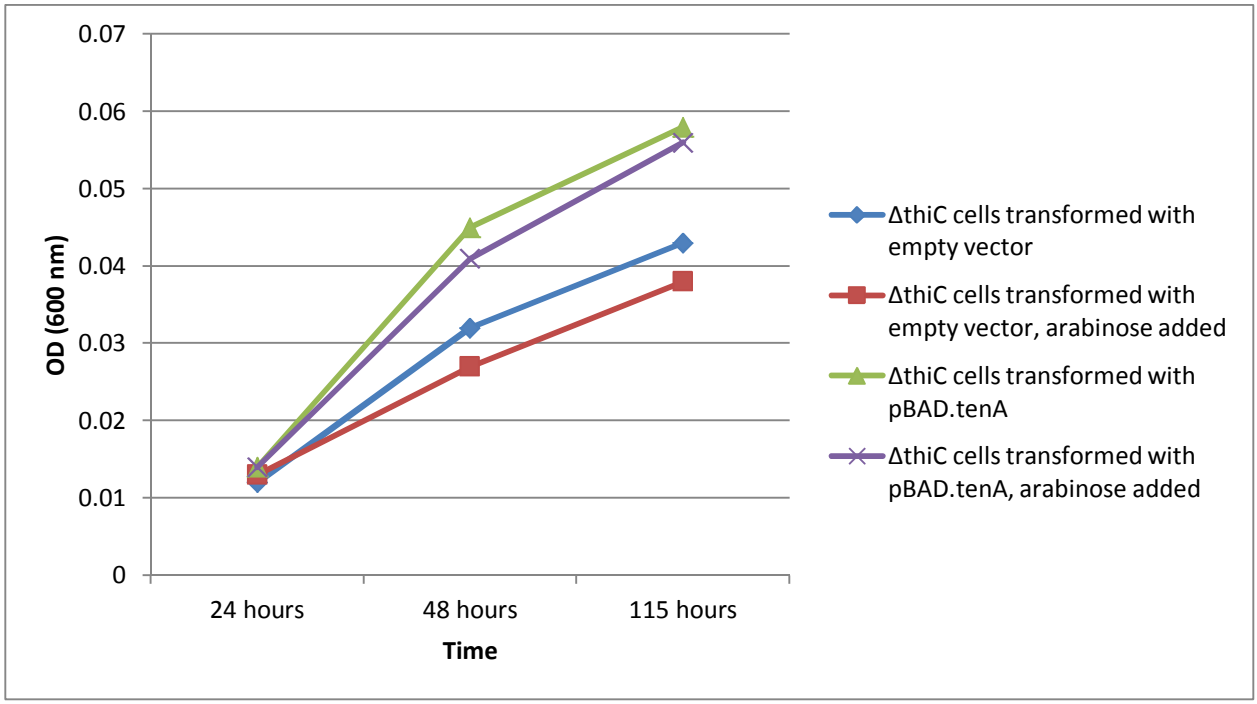


Figure 3: Growth analysis of Δ*thiC* mutant cells in M9 media. This experiment was performed in order to assess the difference in growth rates of mutant *E. coli* cells in M9 minimal media and whether *E. coli thiC* gene is functionally equivalent to SCN's *tenA* gene. Four tubes of mutant *E. coli* cells transformed with either empty pBAD vector or pBAD.LIC.A8.tenA, with or without arabinose were prepared. Arabinose solution was added to verify whether overexpression of the plasmid had any effects on the growth rate OD_{600nm} readings were taken at different time points using the spectrophotometer.

3.5 Cloning of *thiM* cDNA and Δ*thiM* complementation experiment

The cDNA of *thiM* was PCR-amplified using iProof master mix and *thiM* primers. The amplicons were run on the gel and gel-purified. The gel-purification was successful and bands of expected size showed up on the gel. Gel-purified *thiM* cDNA was LIC-cloned into gel-purified, linearized pBAD vector to form pBAD.LIC.A8.thiM, which was used to transform chemically-competent TOP10 *E. coli* cells. These transformed cells were plated on selective LB, ampicillin plates. After overnight incubation, some colonies formed on the plates (no colonies grew on the negative control plates). Several colonies were picked and grown overnight in 5 ml LB, ampicillin media. Plasmid DNA extraction was performed and PCR was carried for all these extracted plasmids, using *thiM* primers. The amplicons were run again for verification. The gel showed that LIC-Cloning of *thiM* was successful and all extracted plasmids contained the insert. Plasmid DNA was sent for sequencing to verify the sequence of the construct and the pBAD.LIC.A8.thiM sequence was correct and had no mutations. Empty pBAD vector and pBAD.LIC.A8.thiM were used to transform electrocompetent Δ*thiM* mutant *E. coli* cells. These transformed cells were selected on LB, ampicillin, kanamycin plates.

For the complementation experiment, the following eight plates were prepared:

1) M9 plate, Δ*thiM* cells with empty vector
2) M9 plate, Δ*thiM* cells with empty vector + arabinose
3) M9 plate, Δ*thiM* cells with pBAD.LIC.A8.thiM
4) M9 plate, Δ*thiM* cells with pBAD.LIC.A8.thiM + arabinose
5) M9 + Thiamine plate, Δ*thiM* cells with empty vector
6) M9 + Thiamine plate, Δ*thiM* cells with empty vector + arabinose
7) M9 + Thiamine plate, Δ*thiM* cells with pBAD.LIC.A8.thiM
8) M9 + Thiamine plate, Δ*thiM* cells with pBAD.LIC.A8.thiM + arabinose

The plates were incubated at 28˚C overnight. From the appearance of the plates it seemed that there was no evidence of complementation. The experiment was repeated using high-purity agar. It was observed that the high-purify agar was softer and initially did not solidify as easily. The pH had to be adjusted in order to make the plates solidify. However, there still did not appear to be any sign of complementation taking place.

Untransformed Δ*thiM* cells were grown in LB and M9 media to compare the difference in growth rates. It was observed that Δ*thiM* grew faster than the other mutant cell lines and reach saturation quickly. Therefore, the transformed cells were inoculated into M9 media and OD_{600nm} readings were taken after 2 hours and 5 hours (to see if any difference could be detected at earlier time points. Figure 4 displays the data from the *thiM* growth curve experiment.

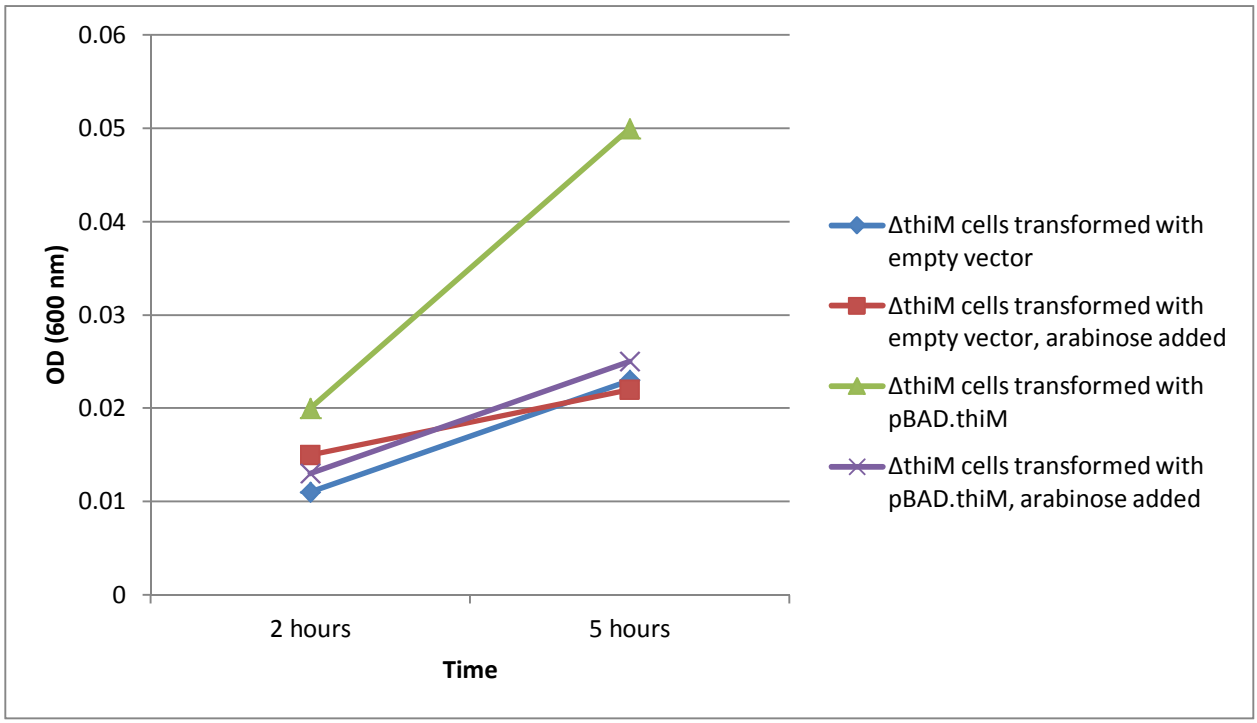


Figure 4: Growth curve analysis of Δ*thiM* mutant cells in M9 media. Four tubes containing 5 ml M9 media were prepared. Two of these tubes were inoculated with mutant cells transformed with pBAD.LIC.A8.thiM and two tubes were inoculated with a loop of mutant cells transformed with empty pBAD.LIC.A8 vector. 50µl of 2% arabinose solution (arabinose acts as an inducer that increases gene expression) was also added to one tube of each type in order to verify whether overexpression of the plasmid had any effects on the growth rate. The cell cultures were incubated at 28ºC in a shaking incubator and OD_{600nm} readings were taken at different time points using the spectrophotometer.

When the growth curve experiment was conducted and readings were taken at 24-hour intervals, there was not a difference in growth between the cells transformed with pBAD.LIC.A8.thiM and the cells transformed with empty vector. Earlier it was observed that untransformed mutant *E. coli* strains that were ordered from the Coli Genetic Stock Center had different growth rates in M9 versus LB media. The Δ*thiM* strain was one strain that had a faster growth rate in both media compared to other strains. Therefore, when the growth curve experiment was repeated, readings were taken after 2 hours and after 5 hours. From the graph in Figure 4, we can see that there was

a difference between the growth rates of cells transformed with pBAD.LIC.A8.thiM and cells transformed with empty pBAD vector. However, when arabinose was present in the media, the growth rate of cells with pBAD.LIC.A8.thiM was similar to the cells transformed with empty pBAD vector. This provided some evidence of complementation. However, *thiM* is part of the thiamine salvage pathway and further experiments would shed light on whether SCN's *thiM* is successful in salvaging thiamine and complementing its corresponding *E. coli* mutant.

3.6 Cloning of *thi4* cDNA and Δ*thiG* complementation experiment

The *thi4* cDNA was PCR-amplified using iProof master mix and *thi4* primers. The amplicons were run on the gel and gel-purified. The gel-purification was successful and bands of expected size showed up on the gel. Gel-purified *thi4* cDNA was LIC-cloned into gel-purified, linearized pBAD vector to form pBAD.LIC.A8.thi4, which was used to transform chemically-competent TOP10 *E. coli* cells. These transformed cells were plated on selective LB, Amp plates. LIC-cloning was successful and colonies formed on the plates (no colonies grew on the negative control plates). Several colonies were picked and grown overnight in 5 ml LB, ampicillin media. Plasmid DNA extraction was performed and PCR was carried for all these extracted plasmids, using *thi4* primers. The amplicons were run again for verification. The gel showed that LIC-Cloning of *thi4* was successful and all extracted plasmids contained the insert. Plasmid DNA was sent for sequencing to verify the sequence of the construct and the pBAD.LIC.A8.thi4 sequence was correct and had no mutations. Empty pBAD vector and pBAD.LIC.A8.tenA were used to transform electrocompetent Δ*thiG* mutant *E. coli* cells. These transformed cells were selected on LB, ampicillin, kanamycin plates.

To check for complementation, the growth curve experiment in M9 minimal media was performed. The results are shown in Figure 5. From this figure we observe that the mutant Δ*thiG* cells transformed with pBAD.LIC.A8.thi4 grew much more compared to Δ*thiG* cells that were transformed with empty pBAD vector. However, when the experiment was repeated, all cell lines showed similar OD readings and similar growth rates. Therefore, the results were inconclusive.

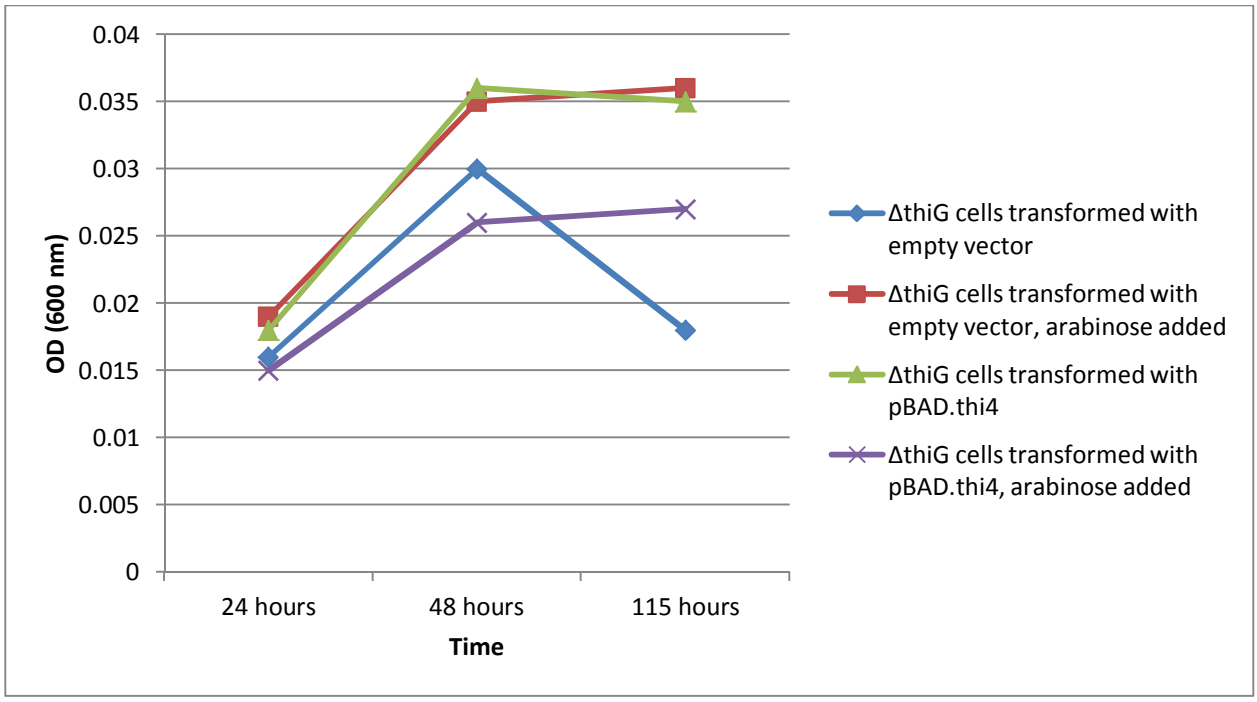


Figure 5: Growth analysis of Δ*thiG* mutant cells in M9 media. This experiment was performed in order to assess the difference in growth rates of mutant *E. coli* cells in M9 minimal media and to test whether *E. coli thiG* gene is functionally equivalent to SCN's *thi4* gene. Four tubes of mutant *E. coli* cells transformed with either empty pBAD vector or pBAD.LIC.A8.thi4, with or without arabinose, were prepared. OD_{600nm} readings were taken at different time points using the spectrophotometer.

3.7 Cloning of *thiE* cDNA and Δ*thiE* complementation experiment

thiE cDNA had to be PCR-amplified several times because after running the PCR-amplicons on the gel, very faint bands or no bands were obtained. When faint bands were obtained, they were cut and gel-purified. When the gel-purified DNA was run again on the gel for verification, no bands were visible. During the process of gel-purification, some DNA always gets lost. It is suspected that there was not that much DNA to begin with and so after the gel-purification step, no *thiE* bands appeared on the gel.

On one or two occasions when faint bands were obtained after gel-purification, the next step, which is LIC-cloning, was performed. However, LIC-cloning was not successful. A few colonies were obtained on the selective plates. These were grown and plasmid DNA was extracted from them and sent for sequencing. But the sequencing results showed that the insert sequence did not resemble *thiE* sequence. Therefore, the cloning step was unsuccessful. It was suspected that since the insert was not in a high enough concentration, cloning efficiency was hampered.

Another reason for failure could be that due too many A's and T's in the *thiE* sequence, the primers we designed were not optimum and the correct *thiE* cDNA was not getting amplified properly. Therefore, artificially-synthesized *HgthiE* cDNA was ordered (GeneArt Strings DNA Fragments, Invitrogen, Life Technologies). Moreover, the following set of primers was ordered which used sequences only from the LIC portion of the insert, not the gene sequence itself.

- *thiE* F primer: 5' TTT AAG ATA TAG ATC ATG C 3'
- *thiE* R primer: 5' TTA TGG AGT TGG GAT CTT ATT ATT A 3'

The artificially synthesized *HgthiE* was cloned into gel purified cut pBAD vector to form pBAD.LIC.A8.thiE. Eight colonies were picked from the LIC-cloning plates, grown in overnight LB media, and plasmid DNA was extracted from them. Plasmid DNA was diluted (1:100) and PCR-amplified using the new set of primers. Plasmid DNA is highly concentrated DNA and it was found that PCR works better if the template DNA is diluted. Too high concentration of template probably increases chance of mis-priming. The amplicons were run on a gel and most of them showed bands of the correct size. The extracted plasmid DNA samples were sent for sequencing to confirm that the *thiE* insert was correct and no mutations arose during the process of cloning. The sequencing results confirmed that indeed cloning was successful and there were no mutations in the *thiE* insert.

The first growth curve experiment data showed that all the cell lines have comparable OD readings at 24-hour intervals. When the growth curve experiment was repeated, the OD readings were again not very different between the cell lines. However, it was interesting to find that when readings were taken after two weeks, it showed that the cell line transformed with empty pBAD vector grew less than the one transformed with pBAD.LIC.A8.thiE (as shown in Figure 6). When the experiment was repeated again and readings were taken after 10 days and 5 weeks, similar results were obtained. These data provided evidence that *HgthiE* was able to complement Δ*thiE* mutant *E. coli* but the complementation became apparent only after a longer incubation period.

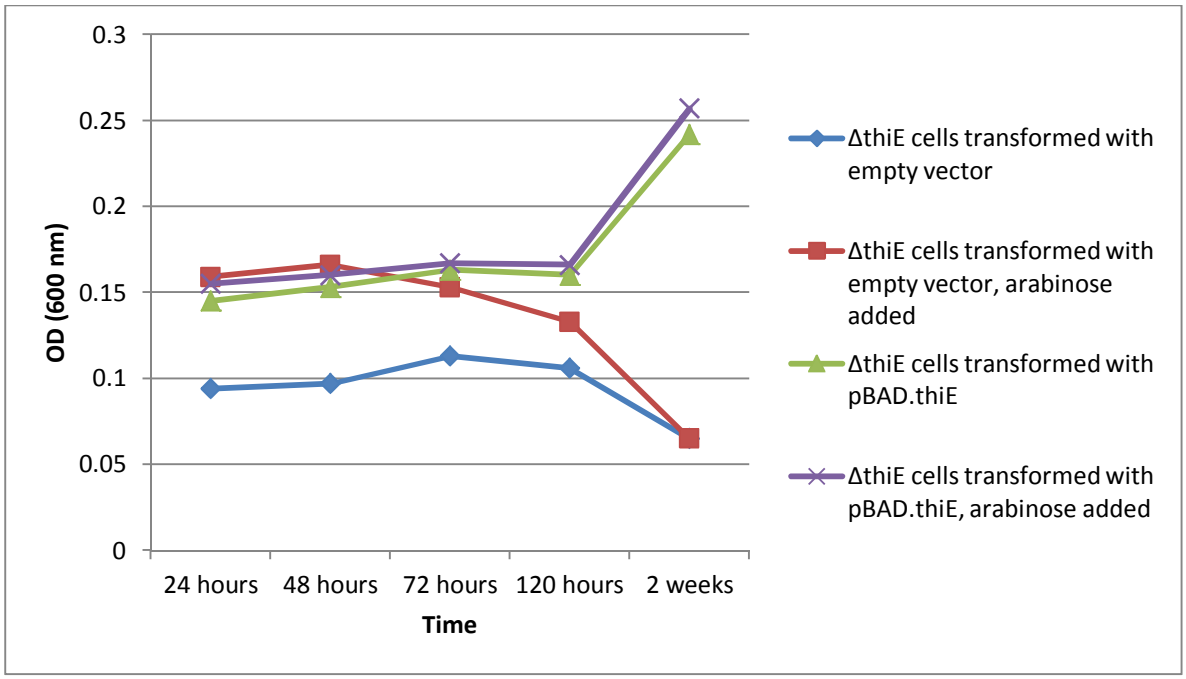
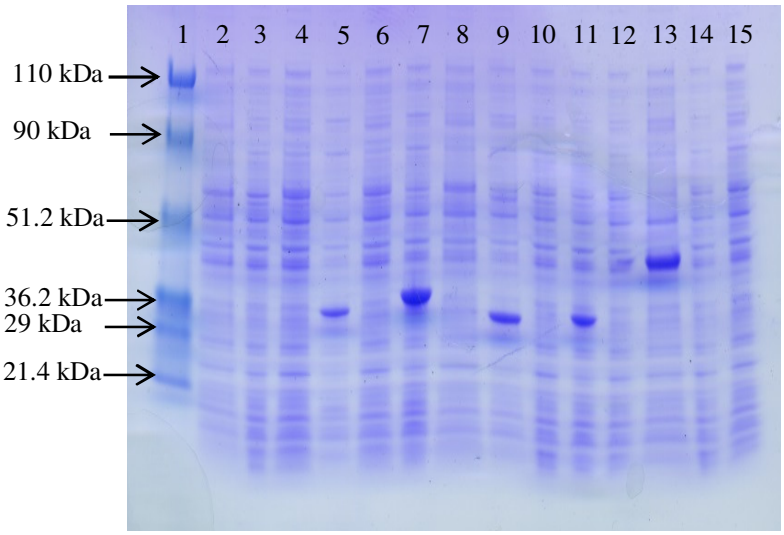


Figure 6: Growth curve analysis of Δ*thiE* mutant cells in M9 media. Four tubes containing 5 ml M9 media were prepared. Two of these tubes were inoculated with mutant cells transformed with pBAD.LIC.A8.thiE and two tubes were inoculated with a loop of mutant cells transformed with empty pBAD.LIC.A8 vector. 50µl of 2% arabinose solution (arabinose acts as an inducer that increases gene expression) was also added to one tube of each type in order to verify whether overexpression of the plasmid had any effects on the growth rate. The cell cultures were incubated at 28ºC in a shaking incubator and OD_{600nm} readings were taken at different time points using the spectrophotometer.

3.8 SDS-PAGE

The SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) showed bands in the *thiD*, *thi4*, *tenA*, *thiM* and *thiE* lanes but no bands for *panC* and *bioB* (i.e., lanes 2, 15). This does not necessarily imply that no *bioB* protein and no *panC* protein are synthesized. Earlier data showed that *panC* was successful in complementing the corresponding *E. coli* mutant. One explanation for this could be that a very small amount of protein is being made which is not enough to show up on the SDS-PAGE gel. Vitamins can be effective in small concentrations and need not be made in copious amounts. High concentrations of biotin and pantothenate can be toxic for the cell and can lead to cell death. Perhaps the only bacterial cells that survived were making small amounts of these proteins which were not visible on the gel.



1. Protein Marker
2. Δ*bioB* cells with empty pBAD vector
3. Δ*bioB* cells with pBAD.bioB
4. Δ*thiM* with empty pBAD vector
5. Δ*thiM* with pBAD.thiM
6. Δ*thiD* with empty pBAD vector
7. Δ*thiD* with pBAD.thiD
8. Δ*thiE* with empty pBAD vector
9. Δ*thiE* with pBAD.thiE
10. Δ*thiC* with empty pBAD vector
11. Δ*thiC* with pBAD.tenA
12. Δ*thiG* with empty pBAD vector
13. Δ*thiG* with pBAD.thiG
14. Δ*panC* with empty pBAD vector
15. Δ*panC* with pBAD.panC

Figure 7: SDS-PAGE gel picture. Arabinose-induced, M9 cell cultures of transformed mutant *E. coli* cells (transformed either with the empty pBAD plasmid or pBAD plasmid containing one vitamin gene of interest) were pelleted and treated with SDS-buffer, which denatures and linearizes the proteins and imparts a uniform negative charge to them. The treated samples were loaded on a 4-20% precast gel and the gel was run at 200 volts for 30 minutes. Electrophoresis separated the proteins according to their size. The gel was stained using the rapid Coomassie staining procedure. The proteins were of the expected sizes.

Table 1: Relevant information about the seven genes studied

Name of gene	Predicted function	NCBI accession number	Size of cDNA (bp)	Predicted size of protein (kDa)	Observed size of protein (kDa)
thiE	Thiamine-phosphate-diphosphorylase	GQ890687	669	24.753	33
thiD	HMP-P kinase/HMP-PP kinase	GQ890688	771	28.527	38
thiM	Hydroxyethylthiazole kinase	GQ890693	879	32.523	36
thi4	Thiazole synthase	GQ890691	972	35.964	45
tenA	Thiaminase II	GQ890692	707	26.159	37
panC	Pantoate β-alanine ligase	GQ890689	1011	37.407	-
bioB	Biotin synthase	GQ890690	1065	39.405	-

The sizes of the proteins (Column 5, Table 1) were estimated from size of cDNAs using a DNA to Protein converter tool available at http://molbiol.edu.ru/eng/scripts/01_06.html.

3.9 Discussion of results

The first time when cloning of *bioB* cDNA was attempted, some mutations were introduced in the insert. This may explain why we did not get full complementation. Mutation can lead to alteration of protein conformation and function. Perhaps the BioB protein made had a different conformation, which hampered its enzymatic activity. The sequence of *bioB* insert turned out to be correct in the second cloning attempt. Several colonies grew in the plates where the mutant cells had been transformed with pBAD.LIC.A8.bioB but it took a long time for these colonies to grow. Therefore, complementation happened but it happened slowly, over the course of two weeks.

When the Δ*panC* complementation experiment was done for the first time, there was some growth in the M9 plate with Δ*panC* cells transformed with pBAD.LIC.A8.panC but it did not look much different from the negative control plate (M9, Δ*panC* cells transformed with empty pBAD vector). We suspected that perhaps some traces of pantothenate from the LB overnight culture allowed cells to grow on the negative control plate and obscured any sign of partial complementation. Therefore, the complementation experiment was repeated and this time the overnight cultures of cells were pelleted and washed several times, to remove any trace of pantothenate that may be present. This additional step proved to be fruitful because the second time the complementation experiment was done, there was no growth in the M9 plate with Δ*panC* cells with empty pBAD vector, and some growth in the M9 plate Δ*panC* cells with pBAD.LIC.panC and M9 + arabinose plate with Δ*panC* cells transformed with pBAD.LIC.panC. This showed that there was some sign of complementation taking place. The growth curve analysis confirmed that pBAD.LIC.panC was able to complement Δ*panC* cells and allow them to grow on M9 minimal media plates.

The results were much stronger for the *thiD* complementation experiment. There was no growth in the M9, empty vector plates and a lot of growth in the M9 plates that had Δ*thiD* cells transformed with pBAD.LIC.thiD. The cells grew just as vigorously as the positive control plates

(M9 + 120µM thiamine). The growth curves were also clearly different for cells that were transformed with empty pBAD vector versus cells that were transformed with pBAD.LIC.thiD.

However, the results of the *thi4* and *thiM* growth curve experiment were not as promising. When the *thiM* growth curve experiment was performed the first time, a difference was seen between the cell lines. However, when the experiment was repeated, all the OD readings were not very similar. Therefore, it was not clear whether complementation was successful. In the first *thi4* growth curve experiment, the OD readings of Δ*thiG* mutant cells transformed with pBAD.LIC.A8.thi4 were slightly greater than the mutant cells transformed with empty pBAD vector. The experiment was repeated and there were no differences between the cell lines. When the experiment was repeated for the third time, the growth rate of cells transformed with pBAD.thi4 was greater than the negative control. However, mutant cells transformed with empty pBAD and induced with arabinose grew at the same rate as the cells transformed with pBAD.thi4. Therefore, it is not clear whether SCN *thi4* were successful in complementing its corresponding *E. coli* mutant.

In case of the *thiE* complementation experiment, initially the results did not look very promising when the OD readings at 2-hour intervals and at 24-hour intervals, the OD readings were not very different between the different Δ*thiE* cell lines. However, when OD readings were taken after three weeks, the cell line transformed with pBAD.LIC.A8.thiE showed an OD reading that was much greater than the mutant cell line transformed with empty pBAD vector. Therefore, it seems that complementation was partial and like the *bioB* experiment, it was slow. Perhaps the complementation was successful only in a subset of cells, and these cells grew slowly over time. For the other thiamine genes we tested, the results were not so clear-cut. Since we are expressing eukaryotic genes in a prokaryotic system, perhaps the different cellular environment inside the *E. coli* cell is affecting protein folding or some other cellular process that is not allowing the protein to attain the right conformation and shape. Therefore, the enzymatic activity of the protein is probably getting affected and the protein is not able to function. We checked whether the protein is being made at all by conducting SDS-PAGE. All the transformed mutant *E. coli* cell lines were induced with arabinose and grown overnight. The SDS-PAGE results showed that all the thiamine biosynthesis genes were expressed and proteins were made. However, there were no protein bands for BioB or PanC. This does not necessarily indicate that no BioB and PanC were

produced. Perhaps small amounts of BioB and PanC proteins were produced, which could not be detected. In the future, silver staining, which is a much more sensitive method, can be used to verify whether BioB and PanC proteins are made in the mutant *E. coli* cells.

3.10 Materials and methods

3.10.1 PCR amplification of cDNA

The following primers were designed such that they contained Ligation Independent Cloning-tags/LIC-tags (LIC tags are underlined)**.** The primers were ordered from Invitrogen, Life Technologies.

- Forward primer for *thiM*: 5'tttaagaaggagatatagatc-ATGtcttacacacctcgcca3'
- Reverse primer for *thiM*: 5'ttatggagttgggatcttatta-attaaaaatagaaagtctcaattcaaaat3'
- Forward primer for *thiE*: 5'tttaagaaggagatatagatc-ATGcttccaaaagacctttttttcac3'
- Reverse primer for *thiE*: 5'ttatggagttgggatcttatta-attattagccataccaaaaaagtaaaacg3'
- Forward primer for *thiD*: 5'tttaagaaggagatatagatc-ATGtcctctcatcgtgc3'
- Reverse primer for *thiD*: 5'ttatggagttgggatcttatta-actgtagccaatctgttgttcaac3'
- Forward primer for *tenA*: 5'tttaagaaggagatatagatc-ATGgaaaaaagtgccaac3'
- Reverse primer for *tenA*: 5'ttatggagttgggatcttatta-aatttcccattggcacatt3'
- Forward primer for *thi4*: 5'tttaagaaggagatatagatc-ATG ctg tcc aaa ctg aag ctg3'
- Reverse primer for *thi4*: 5'ttatggagttgggatcttatta-gacaaggttcagcttgaaacg3'
- Forward primer of *panC*: 5' tttaagaaggagatatagatc-ATGtcgacctcttcaaagttgac3'
- Reverse primer of *panC*: 5' ttatggagttgggatcttatta-ttatccactgataacgtgaaggtac3'
- Forward primer of *bioB*: 5'tttaagaaggagatatagatcatgccacccccacctggctca3'
- Reverse primer for *bioB*: 5'ttatggagttgggatcttattatcacaagttcaaggtcacttttttcgtc3'

PCR conditions: 98°C for 30 seconds; 98°C for 5 seconds; 60°C for 30 seconds, 72°C for 1 minute; 30 cycles; 72°C for 10 minutes.

PCR amplification was conducted using the iProof High Fidelity Master mix. The iProof DNA polymerase was used to reduce the chance of mutation occurring and to reduce the chance of cloning failures.

3.10.2 Ligation Independent Cloning (LIC) of SCN cDNA

When designing the gene-specific primers, Ligation Independent Cloning (LIC) tags were added to the 5' end of the primers in order to create 5' overhangs. A low copy pBAD.LIC.A8 plasmid was chosen as the vector. The pBAD LIC cloning vector was linearized first by treating it with EcoRV restriction enzyme. Both the insert and the vector were PCR-amplified and then gel-purified. For the LIC reaction, when treating the insert and vector with T4 polymerase, the vector was mixed with dCTP whereas the insert was mixed with dGTP. The digestion reaction was conducted for 30 minutes at 22°C followed by an incubation period of 20 minutes at 75°C to denature the T4 DNA polymerase. Once the reaction was complete, the "LICed" insert (i.e., insert that has been treated with T4 polymerase) and the "LICed" vector (i.e., vector that has been treated with T4 polymerase) were combined and incubated at room temperature for 5 minutes to allow annealing of the insert and the vector. (Protocol steps are described in more details in http://www.addgene.org/37501/).

3.10.3 Transformation of chemically competent TOP10 E. coli cells

For the transformation step, 4µl of the combined LICed insert and LICed vector were pipetted into the chemically competent TOP10 *E. coli* cells (Invitrogen, California) and stirred gently. For the negative control, 4µl of the LICed vector alone was added to the cells. The cells were incubated on ice for 30 minutes, heat shocked at 42°C for 30 seconds and immediately transferred to ice again. To allow the cells to recover, 250µl of SOC (Super Optimal broth with Catabolite repression) medium was immediately added to each tube and then the tubes were incubated in the shaking incubator at 37°C for 1 hr. Finally, 50µl-100µl of the transformed cells was spread on to each selective Luria-Bertani (LB) plate with 100µg/µl ampicillin.

3.10.4 Plasmid DNA extraction

Plasmid DNA was extracted using the Omega Plasmid Mini Kit I. Individual colonies from LIC cloning experiment were picked and grown overnight in LB, 100µg/µl ampicillin. Plasmid DNA was then extracted from these overnight (12 to 16 hours) cell cultures. Plasmids were PCR-amplified using gene-specific primers, and then the PCR products were run on 1% agarose gel. This was done to verify whether the extracted plasmids had the desired insert in them. Plasmid

DNA concentration was determined using the spectrophotometer, and then sent for sequencing to the University of Illinois Biotechnology Center.

For all the transformed plasmids constructed, the following sequencing primers were used:

- ARA-F: 5' CATTCTGTAACAAAGCGGGA 3'
- ARA-R: 5' CTGTTTTATCAGACCGCTTC 3'

3.10.5 Preparation of electrocompetent mutant cells

In-frame, single gene knockout, kamamycin-resistant *E. coli* strains from the K-12 Keio collection (Baba *et al.*, 2006) were ordered from the Coli Genetic Stock Center, Yale University. To make these cells electrocompetent, the cells were grown overnight in a 5ml LB broth media solution containing 30 µg/µl kanamycin, in the shaking incubator at 37°C. After 12-16 hours of incubation, the cells were transferred to 500 ml LB media solution containing kanamycin, and grown again in the shaking incubator. The OD_{600nm} (optical density readings at 600 nm) were taken periodically. The cells were harvested when OD_{600nm} reached a value between 0.5 to 1.0. The cells were harvested by chilling the flask for 30 minutes, and then centrifuging in a cold rotor for 15 minutes. The supernatant is removed and the pellet resuspended in 500 ml sterile cold water. The cells are centrifuged again, and the step was repeated but with half the volume of cold water (i.e., 250 ml). In the third step, the cells are resuspended in 10 ml 10% glycerol. Finally, the cells are resuspended in 800µl 10% glycerol. This final cell suspension was then aliquoted in 1.5 ml tubes, frozen in liquid nitrogen and stored at -80°C. The centrifuge bottles, tubes, sterile water and 10% glycerol solution are chilled on ice for several hours before starting the process.

3.10.6 Transformation of mutant E. coli cells

Mutant *E. coli* cells are electroporated by applying an electric current of 1800 volts for 5 miliseconds, and then allowed to recover in SOC medium for an hour. After recovery, the transformed cells are plated in selective LB plates containing 100µg/µl ampicillin and 30 µg/µl kanamycin.

3.10.7 Complementation of single-gene knockout E. coli mutant strains

The transformed mutant *E. coli* strains were plated on four minimal media M9 plates, and the plates are incubated in a shaking incubator at 28˚C. In addition, four M9 + 2% arabinose plates were included in the experiment to see how the inducer arabinose affects growth of the mutants in M9 plates. The negative control plates are plated with mutant cells transformed with the empty pBAD vector. For the *bioB* complementation experiment, avidin, a potent sequestering agent for biotin, was added to the media when making the negative control plates.

3.10.8 Growth curve experiments

The transformed mutant strains were grown on liquid M9 media and monitored over time. Light absorbance readings or optical density readings (OD readings) at 600 nm were taken, using the spectrophotometer.

3.10.9 SDS-PAGE

The transformed mutant cell lines (transformed either with the empty pBAD plasmid or pBAD plasmid containing one of our vitamin genes) were grown for 12-16 hours in LB media, containing ampicillin and kanamycin. 100µl of this culture culture was transferred to fresh tubes containing 5 ml M9 media and the antibiotics. In addition, 2% arabinose (which acts as an inducer and induces gene expression and protein synthesis) was added to each M9 culture tube. After allowing the cells to grow for four hours, 175µl of the cell culture were pipetted into 0.2 ml PCR tubes and pelleted down. Next, 25µl of SDS Buffer was added to each of the pellets and mixed. The PCR tubes were incubated at 100°C for 5 minutes. Using a 4-20%, 15-comb, Mini-Protean TGX precast gel from the Bio-Rad Laboratories, SDS-PAGE was carried out for these treated samples. The gel was run at 200 volts for 30 minutes. The gel was stained using the rapid Coomassie staining protocol. First, isopropanol fixing solution (25% isopropanol, 10% acetic acid) was added to the gel and left undisturbed for an hour. The fixing solution was poured out and rapid Coomassie staining solution (0.006% Coomassie Brilliant Blue G-250, 10% acetic acid) was added. The gel was allowed to shake gently for 16 hours. The staining solution was poured out and the gel was destained with 10% acetic acid solution.

4 Discussion and Future Works

Although vitamins are essential for life, it is widely known that animals have lost their *de novo* vitamin biosynthesis abilities because they can simply absorb them from their diet. It seems paradoxical that a highly-adaptive parasite like SCN, which has the amazing ability to program host cells into a forming a syncytium, is not able to obtain vitamins from the host plant and needs to synthesize its own vitamins. The loss of *de novo* biosynthesis of vitamins in most animals including *C. elegans* suggests that SCN probably also lost the ability to make vitamins, but later reacquired these genes. Although these genes were part of the SCN genome, they were similar to bacterial and fungal genes, which led to the idea that these genes have probably been horizontally transferred from prokaryotes (Craig *et al.*, 2008, 2009). The surprising discovery of these vitamin B biosynthetic genes sparks many questions regarding what roles these vitamins are playing.

Since these vitamins are non-secreted proteins, it is reasonable to think that they are probably not involved in bringing about changes in the host plant and creating of syncytia. Rather, they are probably important to the nematode simply because they are indispensable cofactors that enable the worm to survive. Being parasite and an animal, SCN should have no trouble in getting vitamins from its host plant. However, presence of these vitamin B biosynthesis genes is indicating that the nematode is not getting access to these vitamins from its diet. There could be several explanations for this. Perhaps the nematode is not getting enough vitamins from the host plant simply because of the way it feeds. The vitamins are cofactors and participate in a lot of cellular reactions. Therefore, it seems reasonable to speculate that the vitamins are bound to other proteins and are not freely floating in the plant cytoplasm. SCN is a filter-feeder that uses its stylet to suck in intracellular contents from the syncytium. However, the lumen of the nematode stylet is only 0.1µm (Endo and Wergin, 1988) which is most likely too narrow for the large complexes of vitamins and proteins to pass through.

Another hypothesis is that the plant is shutting down vitamin biosynthesis or actively scavenging vitamins as a defense strategy (Craig *et al.*, 2009). Nutrient restriction, manifested as reduced appetite in animals and humans, is known to be a defensive response to infection (LeGrand, 2000). Nutrient restriction or self-starvation promotes apoptosis of infected cells. Iron-

withholding is a well-studied example of nutrient restriction and it is a common defensive strategy against infection in both plants and animals (Weinberg, 1984). It is known that in the chicken egg, biotin-withholding by avidin is a well-known defense mechanism that is employed to combat microbial invasion. If the host plant is reducing biosynthesis of vitamins as a defense strategy, perhaps SCN has acquired and retained genes for vitamin B biosynthesis in order to evade host defense responses.

It is interesting to note that the thiamine biosynthesis pathway is complete in SCN whereas biotin and pantothenate pathways are partial. One explanation for this could be that the nematode specifically needs thiamine during the earlier developmental stages when the egg develops into a J1 and then molts to a J2. These are times when the nematode has no access to the plant's resources because until a J2 the nematode is successfully able to locate a host plant and establish a feeding site, it has no access to the plant's nutrients, hence no access to the precursors of thiamine-diphosphate. Hence, the nematode is equipped with the complete thiamine biosynthesis pathway to ensure *de novo* biosynthesis of thiamine and survival of the juvenile nematode during these early stages. On the other hand, biotin and pantothenate are probably required in the later stages, when the nematode can obtain necessary precursors from the plant. Therefore, it is sufficient to retain only the gene catalyzing the last step of the pathway in the nematode genome.

Phylogenetic analysis revealed that the *panC* and *bioB* genes were not only present in *H. glycines* but also present in the plant-parasitic root knot nematodes *Meloidogyne incognita* and *Meloidogyne hapla* (Craig *et al.*, 2009). However, *H. glycines* was found to be the only nematode carrying the thiamine biosynthesis genes *thiD*, *thiE*, *tenA*, and *thiM*. In other words, out of the five thiamin biosynthesis genes *thiD*, *thiE*, *tenA*, *thi4* and *thiM*, four of them were found in SCN. However, *thi4* has been found in SCN and in *Pratylenchus vulnus* (*P. vulnus*) a migratory phytoparasite also known as the lesion nematode. The genomics of *P. vulnus* has not been studied and currently there is no information about what thiamine biosynthetic genes are present in the *P. vulnus* genome. The *thi4* gene catalyzes the step of hydroxyethyl thiazole formation, which is a precursor in the pathway and thiamine diphosphate cannot be made with only the *thi4* gene. Therefore, it is likely that *P. vulnus* genome probably also contains the other thiamine biosynthesis genes. It will be interesting if analysis of *P. vulnus* genomic data confirms this.

For future studies, it will be useful to learn whether the host plant represses synthesis of vitamins as a defense mechanism. To test this, GFP fusion proteins to the vitamin B genes can be introduced in plants. The plants can then be infected with SCN. If the fluorescence decreases after the plants are infected with SCN, it would indicate that the plant is down regulating these genes in defensive response to the infection. In addition, experiments should be conducted to find out whether acquisition of these vitamin biosynthesis genes by SCN enables it to evade these defenses. Alternatively, RNA interference (RNAi) (Fire *et al.*, 1998) could be used to post-transcriptionally suppress expression of the vitamin B genes in the nematode (Urwin *et al.*, 2002) so that vitamin biosynthesis is prevented. It will be important to see whether this would hinder the nematode's survival.

5 Conclusion

In this project, complementation studies are performed in order to determine the functionality of seven genes that are part of vitamin biosynthesis/salvage pathways in the soybean cyst nematode. In summary, the cDNAs of the genes were cloned into a low-copy plasmid vector and the recombinant plasmids were transformed into mutant strains of *E. coli* deficient in these genes. To assess functionality of these genes, the mutant cells were grown in M9 minimal media in order to find out whether presence of these genes rescues the mutant and enables it to grow. Results from these experiments show evidence that these genes are expressed and make soluble proteins and some of them also show that they can complement their corresponding *E. coli* mutants and enable them to grow better in M9 minimal media. Therefore, these genes are functional and they are not pseudogenes. Presence of acquired vitamin B biosynthesis genes in SCN implies that they are providing certain advantages to the nematode. If these SCN genes do play a pivotal role in assisting the nematode to survive within the host plant, then blocking the activities of these genes may be a new approach of controlling this devastating phytoparasite.

References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2: 2006 0008 10.1038/msb4100050.

Begley, T. P., Downs, D. M., Ealick, S. E., McLafferty, F. W., Van Loon, A. P., Taylor, S., Campobasso, N., Chiu, H. J., Kinsland, C., Reddick, J. J., and Xi, J. 1999. Thiamin biosynthesis in prokaryotes. Arch. Microbiol. 171:293-300.

Bekal, S., Niblack, T. L., and Lambert, K. N. 2003. A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. Mol. Plant-Microbe Interact. 16:439-446.

Burrows, P. R., and Stone, A. R. 1985. *Heterodera glycines*. CIH Descriptions of Plant-Parasitic Nematodes No. 118. CAB International, Wallingford, UK.

Craig, J. P. 2009. Analysis and characterization of vitamin B biosynthesis pathways in the phytoparasitic nematode heterodera glycines. (Order No. 3395583, University of Illinois at Urbana-Champaign). ProQuest Dissertations and Theses, , 83. Retrieved from http://search.proquest.com/docview/89185802?accountid=14553. (89185802).

Craig, J. P., Bekal, S., Hudson, M., Domier, L., Niblack, T., and Lambert, K. N. 2008. Analysis of a horizontally transferred pathway involved in vitamin B6 biosynthesis from the soybean cyst nematode *Heterodera glycines*. Mol. Biol. Evol. 25:2085-2098.

Craig, J. P., Bekal, S., Niblack, T., Domier, L., and Lambert, K. N. 2009. Evidence for horizontally transferred genes involved in the biosynthesis of vitamin B1, B5 and B7 in Heterodera *glycines*. J. of Nematol. 41:281-290.

Dewick, P. M. 1998. The biosynthesis of shikimate metabolites. Natural Products Reports 15:17 58.

Doyle, E. A., and Lambert, K. N. 2002. Cloning and characterization of an esophageal-gland-specific pectate lyase from the root-knot nematode *Meloidogyne javanica*. Mol. Plant Microbe Interact. 15:549-556.

Endo, B. Y., and Wergin, W. P. 1988. Ultrastructure of the second-stage juvenile of the root-knot nematode, *Meloidogyne incognita*. Proc. Helminthol. Soc. Wash. 55:286-316.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature, 391, 806–811.

Genschel, U., Powell, C. A., Abell, C., and Smith, A. G. 1999. The final step of pantothenate biosynthesis in higher plants: cloning and characterization of pantothenate synthetase from *Lotus japonicus* and *Oryza sativum* (rice). Biochem. J. 341:669-678.

Grundle, F., Sobczak, M., and Golinowski, W. 2004. Formation of wall openings in root cells of *Arabidopsis thaliana* following infection by the plant-parasitic nematode *Heterodera schachtii*. European Journal of Plant Pathology 104:545-551.

Harris, R. A., Bowker-Kinley, M. M., Huang, B., and Wu, P. 2002. Regulation of the activity of the pyruvate dehydrogenase complex. Adv. Enzyme Regul. 42: 249–259.

Jitrapakdee, S., and Wallace, J. C. 2003. The biotin enzyme family: conserved structural motifs and domain rearrangements. Curr Protein Pept Sci 4:217-229.

Jurgenson, C. T., Begley, T. P., and Ealick, S. E. 2009. The structural and biochemical foundations of thiamine biosynthesis. Annu. Rev. Biochem 78:569–603.

Knowles, J. R. 1989. The mechanism of biotin-depending enzymes. Annu. Rev. Biochem. 58:195-221.

Lambert, K. N., Allen, K. D., and Sussex, I. M. 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode, *Meloidogyne javanica*. Mol Plant Microbe Interact. 12:328-336.

Lambert, K. N., Bekal, S., Domier, L. L., Niblack, T. L., Noel, G. R., and Smyth, C. A. 2005. Selection of *Heterodera glycines* chorismate mutase-1 alleles on nematode-resistant soybean. Mol Plant Microbe Interact 18:593-601.

LeGrand, E. K. 2000. Why infection-induced anorexia? The case for enhanced apoptosis of infected cells. Medical Hypotheses 54:597-602.

Molinari, S., and Miacola, C. 1997. Antioxidant enzymes in phytoparasitic nematodes. J. of Nematol. 29:153-159.

Niblack, T. L., Arelli, P. R., Noel, G. R., Opperman, C. H., Orf, J. H., Schmitt, D. P., Shannon, J.G., and Tylka, G. L. 2002. A revised classification scheme for genetically diverse populations of *Heterodera glycines*. J. Nematol. 34:279-288.

Niblack, T. L., Lambert, K. N., and Tylka, G., L. 2006. A model plant pathogen from the Kingdom Animalia: *Heterodera glycines*, the soybean cyst nematode. Annu. Rev. Phytopathol. 44: 283-303.

Park, J. H., Dorrestein, P. C., Zhai, H., Kinsland, C., McLafferty, F. W., and Begley, T. P. 2003. Biosynthesis of the thiazole moiety of thiamine pyrophosphate (Vitamin B1). Biochem. 42: 12430-12438.

Qin, L., Kudla, U., Roze, E. H., Goverse, A., Popeijus, H., Nieuwland, J., Overmars, H., Jones, J. T., Schots, A., Smant, G., Bakker, J., and Helder, J. 2004. Plant degradation: a nematode expansin acting on plants. Nature 427:30.

Riggs, R. D. 1982. Cyst nematodes in the Southern USA. In: Nematology in the Southern Region of the United States (Ed. by Riggs, R. D.) Cooperative Series Research Bulletin 276:77-95.

Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. D., Krell, T., Coggins, J. R., Coombs, G. H., Milhous, W. K., Tripori, S., Ferguson, D. J. P., Chakrabarti, D., and McLeod, R. 1998. Evidence for the shikimate pathway in apicomplexan parasites. Nature 393:801-805.

Schmid, J., and Amrhein, N. 1995. Molecular organization of the shikimate pathway in higher plants. Phytochem. 39:737-749.

Schmitt, D. P., Wrather, J. A., and Riggs R. D. 2004. Biology and Management of Soybean Cyst Nematode Second Edition. Marceline, Missouri: Schmitt & Associates of Marceline.

Smant, G., Stokkermans, J. P., Yan, Y., de Boer, J. M., Baum, T. J., Wang, X., Hussey, R. S., Gommers, F. J., Henrissat, B., Davis, E. L., Helder, J., Schots, A., and Bakker, J. 1998. Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc. Natl. Acad. Sci. USA 95:4906-4911.

Streit, W. R., and Entcheva, P. 2003. Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. Appl Microbiol Biotechnol 61: 21-31.

Urwin, P. E., Lilley, C. J., and Atkinson, H. J. 2002. Ingestion of double-stranded RNA by preparasitic juvenile cyst nematodes leads to RNA interference. Mol. Plant–Microbe Interact. 15, 747–752.

Webb, E., and Downs, D. 1997. Characterization of *thiL*, encoding thiamine-monophosphate kinase, in *Salmonella typhimurium*. J. Biol. Chem. 272: 15702-15707.

Webb, M. E., Smith, A. G., and Abell, C. 2004. Biosynthesis of pantothenate. Nat. Prod. Rep 21: 695-721.

Weinberg, E. D. 1984. Iron withholding: A defense against infection and neoplasia. Physiol Rev 64:65-102.

Wrather, J. A., Anderson, T. R., Arsyad, D. M., Tan, Y., Ploper, L. D., Porta-Puglia, A., Ram, H. H., and Yorinori, J. T. 2001. Soybean disease loss estimates for the top ten soybean-producing countries in 1998. Can. J. PlantPathol. 23:115-121.

Wrather, J. A., and Koenning, S. R. 2006. Estimates of disease effects on soybean yields in the United States 2003 to 2005. J. Nematol. 38:173-180.

Wrather, J. A., and Koenning, S. R. 2009. Effects of diseases on soybean yields in the United States 1996 to 2007. Online. Plant Health Progress doi:10.1094/PHP-2009-0401-01-RS. Retrieved from http://www.plantmanagementnetwork.org/pub/php/research/2009/yields/

Wrather, J. A., Stienstra, W. C., and Koenning, S. R. 2001. Soybean disease loss estimates for the United States from 1996 to 1998. Can. J. Plant Pathol. 23:115-121.

Zacheo, G., Arrigoni-Liso, R., Bleve-Zacheo, T., Lamberti, F., and Arrigoni, O. 1983. Mitochondrial peroxidase and superoxide dismutase activities during the infection by *Meloidogyne incognitia* of susceptible and resistant tomato plants. Nematol. Medit. 11:107-114.

Zhang, Y., Taylor, S. V., Chiu, H., and Begley, T. P. 1997. Characterization of the *Bacillus subtilis* thiC operon involved in thiamine biosynthesis. J. Bacteriol. 179:3030-3055.